A Demonstration of a Spoken Dialogue Interface to an Intelligent Procedure Assistant for Astronaut Training and Support

Gregory Aist, John Dowding, Beth Ann Hockey, Jim Hieronymus

Research Institute for Advanced Computer Science
RIACS, Mail Stop T27A-2
NASA Ames Research Center
Moffett Field CA 94035-1000 USA
{aist, jdowding, bahockey, jimh}@riacs.edu

1   Introduction

One of the main goals of the RIALIST group at Ames is to support Human Exploration and Development of Space (HEDS) by applying spoken dialogue technology to training and task performance on the ground and in space. We are moving to address this goal by developing a system aimed at taking technical documentation from the International Space Station and producing spoken dialogues geared towards astronaut training and task support. In this demo abstract, we will introduce the scope and nature of tasks performed on board the International Space Station, describe our system architecture, and discuss the range of functionality afforded by the current system. In the demonstration itself, conference participants will be provided a hands-on experience with our system, which converts a technical procedure written as a set of logical clauses into a spoken dialogue interaction aimed at astronaut training and support. They will be able to use the system to walk them through a training exercise aimed at using a digital camera to take pictures of the Earth from (simulated) outer space.

2   Overview of Astronaut Training

The astronaut training process currently consists of a number of phases (NASA 1995). The first phase of basic study, takes approximately 1 year where the individual reviews a broad spectrum of science, including astronomy, computer science, Earth resources, guidance and navigation, mathematics, meteorology, and physics. This training is accomplished primarily with lectures, briefings, and individual study with workbooks and computer-based training (CBT). Training progresses with Single Station Training (SST), where the astronaut practices the operation of a particular system, usually in a physical mock-up of the control interface for that system. The astronaut practices predefined procedures for both nominal and off-nominal conditions, either solo, or with a training specialist on that system. These first two phases occur before an astronaut is placed on a team.

Once assigned to a team, the astronaut will continue in Phase Training for 7-12 months preparing for a particular flight. This team training is carried out in the Shuttle Mission Simulator, and is directed by a specialized team dedicated to the particulars of a specific flight.

Of course, an astronaut's commitment to training persists throughout his or her career. Furthermore, extended stays in space mean time away from regular training facilities. For longer missions, such as the International Space

Station, or manned missions to Mars, there is an additional interest in training to refresh skills just prior to use. Thus just-in-time automated training (e.g. Yu 1997) may well play a role.

We are developing a spoken dialogue system for training and task support aimed at tasks typically performed on board the International Space Station – with training deployment envisioned initially, but also directed at eventual on orbit deployment. In related work which includes in-task support, Rudnicky et al. (1996) describe supporting vehicle maintenance with speech interfaces; the Mission Rehearsal Exercise (USC n.d.) aims at supporting "mission-oriented training" for military personnel.

3   System Architecture

We developed a procedure assistant consisting of the following components (Figure 1). First, the speech recognizer (Nuance, using specially compiled language models (Dowding et al. ACL 2001)), recognizes the user's speech and produces a string of words which are passed to the Gemini parser (Dowding et al. ACL 1993). Gemini outputs logical forms that are passed to the discourse manager (DM). The discourse manager uses a translation function to change resolved logical forms into actions that the procedure manager (back end) can handle. The procedure manager loads a new procedure into memory, and maintains it in memory as a list structure. The procedure manager also navigates between steps in a procedure, and generates task state reports that are then passed to the dialogue manager. For example, a task state report might be [procedure_loaded(camera), "Take pictures of the Earth."]. The dialogue manager takes the task state report, constructs a combination of logical form (LF) and/or word strings, and passes that to the generator. The generator then constructs strings for the text-to-speech engine (Festival) to say, based on the command received from the dialogue manager. Each
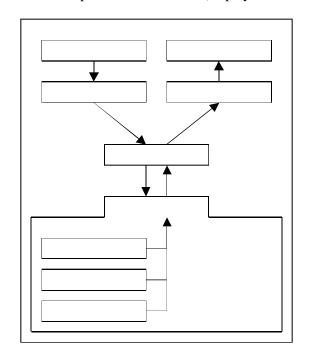
procedure is stored as a separate Prolog file, in order to allow for dynamically switching between procedures.

The procedure is specified as a tree, with each node consisting of a description of the node and a list of its children. The description of the node is written as an unanalyzed imperative. Separate left and right lists of nodes are kept so that the current position in the node's children can be maintained - the head of the right list is always the current position in that list. The procedure is loaded into memory as a Prolog term with the structure of subprocedures left in unexpanded form, and instantiated when needed.

Current procedures available include a procedure to conduct the initial portion of a periodic physical exam, and a procedure to take pictures of the Earth from space.

4   An example procedure:
Photograph the Earth from space

We designed a procedure that walks the user through taking a picture of a view of Earth from space. The procedure consists of unpacking and assembling a digital camera, and then using it to take several pictures of the Earth (displayed on a

nearby computer monitor.)

The intelligent procedure assistant currently includes the following capabilities:

1. Load a procedure into the procedure manager.

2. Say again - repeat a prompt.

3. Next step - move forward in the procedure.

4. Drill down - unpack a procedure into its component parts.

5. Speak up – increase output volume (and its counterpart, decrease output volume).

References

J. Dowding, B. A. Hockey, J. M. Gawron, and C. Culy. 2001. Practical issues in compiling typed unification grammars for speech recognition. ACL 2001.

J. Dowding, M. Gawron, D. Appelt, L. Cherny, R. Moore, and D. Moran. 1993. Gemini: A natural language system for spoken language understanding. ACL 1993.

NASA. 1995. Selection and Training of Astronauts.
http://liftoff.msfc.nasa.gov/academy/astronauts/training.html

Rudnicky, A. I., Reed, S. D., and Thayer, E. H. 1996. SpeechWear: A mobile speech system.
http://www.speech.cs.cmu.edu/air/papers/speechwear.ps

USC. N.d. Mission Rehearsal Exercise.
http://www.ict.usc.edu/misreh.html

Yu, Leo. 1997. Just-In-Time Training.
http://ic.arc.nasa.gov/ic/projects/JITT/